Associating Facebook Measurable Activities with Personality Traits: A Fuzzy Sets Approach

Misirlis, Nikolaos and Lekakos, George and Vlachopoulou, Maro

University of Macedonia, Athens University of Economics and Business, University of Macedonia

9 February 2018

Online at https://mpra.ub.uni-muenchen.de/90100/
MPRA Paper No. 90100, posted 22 Nov 2018 07:29 UTC
Abstract: In this study we identify potential associations between people’s personality (utilizing the popular Big Five personality model) and measurable Facebook activities such as number of likes received, number of posts, number of comments on posts. Extant literature suggests that personality can be manifested through different features of the Facebook profiles but under an implicit assumption that those users may belong in a single psychographic group. However, it has been shown that people may share characteristics, common acts and behaviors of more than one psychographic group. In this study we aim to address limitations of previous studies, by adopting a fuzzy set approach which is capable to handle users’ membership in multiple psychographic groups. Furthermore, fsQCA offers equifinality, which means that research can end up to the same outcome, beginning from different initial combinations of data. The work presented here provides empirical evidence concerning the association between Facebook activities and users’ personalities in a novel way indicating the significance of this relationship and providing alternative combinations that lead to the same output. Furthermore, it paves the ground towards predicting social platforms’ measurements, other than Facebook, relying on users’ personalities, using the same technique but on different fields of study and social media platforms.

Keywords: Big Five; fsQCA; Facebook; personality traits; Facebook measurements.

JEL Classification: L86, C8, L82

Biographical note: Dr. Misirlis Nikolaos is a Software engineer and social media analyst; he currently teaches ICT at American Farm School and American College of Thessaloniki, Greece. George Lekakos is Associate Professor at the Department of Management Science and Technology, Athens University of Economics and Business, Greece, and Director of the MSc in Management Science and Technology. He leads the Intelligent Media Lab (IML) within the ELTRUN Research Center at the Athens University of Economics and Business. Maro Vlachopoulou is former Dean of School of Information Sciences and currently Professor at the Department of Applied Informatics of the University of Macedonia, Greece. Corresponding author: mavla@uom.gr

1 INTRODUCTION

Internet has a major impact on social life with the self-presentation need or the need to belong somewhere to be important (Nadkarni & Hofmann, 2012; Seidman, 2013). Nowadays, human interaction through social media platforms is rapidly increasing, providing researchers the opportunity to understand human behavior exploiting the rich information available in social media (Carpenter, Green, & LaFlam, 2011). Facebook, with over 1.7 billion registered users and 750 million daily logs on, represents the most evolving and popular social networking site among all social media platforms (Ding, Cheng, Duan, & Jin, 2017; Sedghi, 2014). Facebook has become a useful tool not only for daily friendship interaction but also for seeking information, jobs and advertisings. Extant research suggests that there is an association between Facebook activity data and personality, demonstrating a relationship between users’ online and offline activities. This indicates that users’ personality can be extracted from their Facebook activity (Gosling, Augustine, Vazire, Holtzman, & Gaddis, 2011). Facebook profiles reflect actual personality and not just a self-idealization and therefore it is suggested that people use Facebook to communicate their real personality and accurate personal profiles (Amichai-Hamburger & Vinitzky, 2010; Garcia & Sikström, 2014). As a result, it can be claimed that Facebook users do not try to misrepresent their profiles but on the contrary, Facebook profiles represent the true personality of a person. Obviously, fake profiles exist on Facebook, but the current research did not try to exclude these profiles as they
belong to the accepted statistical error. Associating social media and users’ interactions is an old field of study. Katz (1959) asked what do people do with the media studying the uses of gratification theory, explaining that certain media contents have greater appeal than others. The same theory assumes that an audience of a media is not homogeneous. This model has been used in order to study online content on Facebook (Gosling et al., 2011). By predicting the preference of users on a product, advertising companies can personalize advertisements to each individual or each group of individuals with common preferences. With the growth of social networks, social marketing is also growing fast (Yadav, de Valck, Hennig-Thurau, Hoffman, & Spann, 2013; Chatzigeorgiou, 2017). Many customers relay their decision to buy upon reviews of other customers. An individual who is predicted to click more the 'like button' is a better advertiser of a new product instead of an skeptic one. Previous studies suggest that personality can be manifested through different features of the Facebook profiles. Bachrach et al. (2012) provided empirical evidence demonstrating that counting the number of published photos, events and groups and the objects that the user has liked on Facebook we can determine a users’ personality. In this study an opposite procedure is followed. Instead of predicting personalities from simply measured activities, we extract the personality of users based on Big Five theory and associate them with the measurable activities of users on Facebook. For that purpose, a fuzzy set approach is adopted. We build upon extant literature and extend our knowledge by examining specific uses of Facebook associating them with specific personality traits. The aim of this study is to investigate associations between measurable activities on Facebook and users' personality. By using the diverse combinations of fsQCA, we arrive to what we call an equifinal outcome, which means that we obtain the same effect or outcome from initially different events, in our case combinations. Equifinality together with the fact that personalities are combined together represent the coverage of the gap of the research as well. What it is not yet fully analyzed on the literature, to the best of our knowledge, is the fact that personality traits from the Big Five model can be associated with measurable activities with diverse methods. Theories are used to explain the online behavior of users in combination (Thygeson, Peikes, & Zutshi, 2013). This taxonomy is used to study the data with a deeper perspective (El Sawy, Malhotra, Park, & Pavlou, 2010).

The advantages of the fsQCA methods are: a) equifinality and b) for n elements the method produces max 2^n different combinations reducing the causality factor and helping researchers to design more adaptable strategies for their research on Information System field (Cárdenas, 2012; Downey & Stanyer, 2010; Mozas-Moral, Bernal-Jurado, Medina-Viruel, & Fernández-Uclés, 2016; Vis, 2012). By this we mean that alternative sources, elements or individuals can be used every time differently in order to obtain similar outcomes. Furthermore, fsQCA can reveal relationships between variables and outcomes even if the sample is small, as in our case (Thygeson, Peikes, & Zutshi, 2013). Literature is already associating personality with Facebook’s measurable activities with diverse methods. Theories are used to explain the online behavior of users in combination with their personality (Gosling et al., 2011; Hall & Pennington, 2013). One of the most used models to categorize users and personalities is the Big Five model (De Raad, 2000; John, Naumann, & Soto, 2008; Zuckerman, Kuhlman, Joireman, Teta, & Kraft, 1993). This taxonomy is one of the most reliable methods for exporting and monitoring personalities (McCrae & John, 1992; Moore & McElroy, 2012; Ryan & Xenos, 2011; Tan, 2012; Zywica & Danowski, 2008). These five traits are defined as follows (John et al., 2008):

### 2 FSQCA IMPORTANCE AND RESEARCH PROPOSITIONS

FsQCA as a method is used more and more often because of its capacity to study the data with a deeper perspective (Pappas, Kourouthanassis, Giannakos, & Chrisissipoulos, 2015). Especially in social science, it is not always possible for qualitative data to fit in just one category. Fuzzy set procedure solves this limitation by simultaneously dealing data with both qualitative and quantitative way (Zschoch, 2011). Several studies have been conducted regarding fsQCA. Precisely, several studies are related to: behavioral analysis (Pappas et al., 2015), tourism (Wu, 2015), purchase intentions (Gunawan & Huaring, 2015), education (Y.-C. Choi & Lee, 2015; Stevenson, 2013) or business-(Kask & Linton, 2013; Mes-Verdú, Ribeiro-Soriano, & Roig-Tierno, 2015; Yilmaz, Varnali, & Kasnakoglu). In fsQCA in each case is assigned a score in the interval between 0 and 1, expressing the membership of a case in the available groups (i.e. Big Five groups in our research). The basic idea behind fuzzy set is to permit the scaling of membership scores and thus allow partial membership, rather than just the presence (1) or the absence (0) of an element. In our research we use the continuous fuzzy set interval which uses I for fully-in membership and 0.5<rank<1 for a degree of membership which is “more in than out” in a certain group. 0.5 is the crossover point which means neither in nor out. 0<rank<0.5 is for degree of membership which is more out than in and 0 for non-membership. The correspondence to the 0-1 range is achieved through the calibration function of fsQCA. The fsQCA provides the researchers the advantage of viewing the elements in combination rather than separately. Ragin and Fiss (2008) assert that in fsQCA method equifinality is accommodated which means that a different set of combinations can produce the same result. Such configuration theories examine the elements as a set and not separately and are suitable for Information Systems research (El Sawy, Malhotra, Park, & Pavlou, 2010).

The remaining of the paper is organized as follows. In the next section we present the relevant literature for the fsQCA method and its importance. Next, we present how personality traits are associated with the measurable activities of users on Facebook. In order to design more adaptable strategies for their research on Information System field (Cárdenas, 2012; Downey & Stanyer, 2010; Mozas-Moral, Bernal-Jurado, Medina-Viruel, & Fernández-Uclés, 2016; Vis, 2012). By this we mean that alternative sources, elements or individuals can be used every time differently in order to obtain similar outcomes. Furthermore, fsQCA can reveal relationships between variables and outcomes even if the sample is small, as in our case (Thygeson, Peikes, & Zutshi, 2013). Literature is already associating personality with Facebook’s measurable activities with diverse methods. Theories are used to explain the online behavior of users in combination with their personality (Gosling et al., 2011; Hall & Pennington, 2013). One of the most used models to categorize users and personalities is the Big Five model (De Raad, 2000; John, Naumann, & Soto, 2008; Zuckerman, Kuhlman, Joireman, Teta, & Kraft, 1993). This taxonomy is one of the most reliable methods for exporting and monitoring personalities (McCrae & John, 1992; Moore & McElroy, 2012; Ryan & Xenos, 2011; Tan, 2012; Zywica & Danowski, 2008). These five traits are defined as follows (John et al., 2008):
• Openness measures peoples’ originality and openness. (Cukić & Bates, 2014). Open to new experience users are correlated with often status updates and participation to Facebook groups (Bachrach et al., 2012).
• Conscientiousness measures the constraint and the control of impulse. Such impulses are thinking before acting, delaying gratification, following rules and being organized. This exact type of personality implies that conscientious users are hesitant with likes but not with photo uploads (Bachrach et al., 2012).
• Extraversion measures a person’s energy and enthusiasm. Extraverted individuals usually have positive way of thinking (Augustine & Hemenover, 2008) and tend to like posts quite often.
• Agreeableness measures a person’s altruism and affection. Agreeableness may also refer to individuals who seek information on internet (J. Choi & Kim, 2014; Nadkarni & Hofmann, 2012; Seidman, 2013).
• Neuroticism measures a person’s negative emotionalness and nervousness (John et al., 2008; Smith, Saklofske, & Nordstokke, 2014). Neurotic individuals often hide some aspects of themselves, but they show them only online (Seidman, 2013).

In literature there are many studies that suggest relationship between personality and Facebook usage. Some of them indicate which personalities use Facebook under certain conditions (Carpenter et al., 2011; Christou, 2015). Others (Amichai-Hamburger & Vinitzky, 2010; Ross et al., 2009) found that users with high neuroticism have accurate personal profile information or that users with high extraversion use frequently the internet. Further studies shown that high extraverted and open to new experiences users are less influential that it was though on past studies (Correa, Hinsley, et al., 2010) in order measure the psychological profile of each personality types of Big Five.

The goal of this research is to develop a model that associates Facebook users’ personalities with measurable online activities. More specifically, in the following sections we examine if the five personality traits from Big Five model can be used as classifiers (predictors) of three easily measurable types of online activities, namely (a) number of posts (the number of posts a user wrote in a certain period of time), (b) number of likes (users received for their posts in a certain period of time) and (c) comments of posts (the number of comments received on users’ posts in a certain period of time). The above study of personality traits and online activity lead to the next research proposition:

**Proposition 1**: Users with different combinations of personalities may perform the same online activities.

Based on the above proposition, the methodology presented in next section was developed.

3 METHODOLOGY

3.1 Data collection

To recruit respondents, we created a group on Facebook where each member had to compile a 44-item questionnaire (Arterberry, Martens, Cadigan, & Rohrer, 2014; Plaisant et al., 2010) in order measure the psychological profile of each user according to the Big Five model (John, Donahue, & Kentle, 1991). The questionnaire was named in order to associate response with Facebook profiles. To achieve that, an online Facebook application was used, Netvizz (https://apps.facebook.com/netvizz/). Netvizz permits to download the online activities for each user of the group, such as number of likes, number of posts, number of comments etc. Consequently, the data were visualized with Gephi (Bastian, Heymann, & Jacomy, 2009). In a total of 134 users, 80 responded, giving a response rate equal to 59.7%. The participation was voluntary with no economical or other reward for the participants.

3.2 Respondents

The average age of our group of people is 29 years old with 51% of them females and 49% males. 53% of the participants are university graduates, with a master or a PhD. 40% were still university students at the time of the participation. After filling the questionnaire and exporting for all users their personality profile, we collected Facebook data for a period of 7 months. These data represent the activities of our friends in the network. These data were the number of likes they received, the number of posts and the number of comments a user received on posts.

3.3 Survey instruments

The questionnaire consists of 44 statements – short phrases with relatively accessible vocabulary (John et al., 1991; John et al., 2008). Each statement must be rated on how much a user agrees with that statement on a five point scale: (1) disagree, (2) slightly disagree, (3) neutral, (4) slightly agree, and (5) agree. Each question gives a rate in one of the five personality types of Big Five.

3.4 Analysis

As discussed above, the aim of this research is to examine whether the Big Five personality traits can be used as predictors for users’ measurable activities on Facebook. We take account that users do not belong to a single psychographic group since they may share characteristics of more than one psychographic trait. To our knowledge this is the first study, oriented on personality and social media activities correlation that uses fsQCA methodology in order to examine independent variables that lead to an outcome. Qualitative data as such data need to be calibrated, before fuzzy set values, in a range of 0 to 1, can be determined (Basurto & Speer, 2012). There is no clear indication in literature on how the transformation will be achieved. Furthermore, authors follow different strategies in order to calibrate the data. In our case, the fsQCA methodology and the calibration process is applied as follows. The Big Five questionnaire exports results from likert-based questions from 1 to 5, although measurements from Facebook lack of continuity. Value 1 denotes full set membership and 0, no set membership. In order to calibrate the dataset, we must use the calibration function of fsQCA. This function uses 3 values as thresholds (full membership, full no-membership and the crossover point). To define these thresholds we used the boxplot in SPSS 22. Boxplot provides information about how the data is spread out and centered, so extreme values can be excluded. With this method, the three qualitative anchors for the calibration are shown analytically on table 1:
Once the calibration is completed, the truth table is produced. Truth table, then, is refined concerning the consist ency and the frequency of the results. The five categories of the Big Five were used as independent variables so as to examine their impact on the dependent variables (number of likes, posts, number of comments on posts). A lot of discussion exists around the type of solution to select to analyze, among the three provided from the tool. The right solution to choose varies from a series of variables to take into consideration. Parsimonious solution requires having high consistency and high coverage, a quite rare scenario. In our case, each causal path matters due to limited variable used, so a ‘strict’ solution that excludes some of the paths, is not preferred. Furthermore, most of the similar articles on the literature review, confirm that the intermediate solution is the most secure way in order not to exclude any useful path. The present research uses the intermediate solution due to the aforementioned reasons (Baumgartner, 2015; Engeli, Allison, & Allison, 2014) All the possible combinations of personalities are presented to Tables 2 to 4, used as predictors for the three different outcomes (intention to post, prediction to receive more likes on posts, prediction to receive more comments on posts, respectively for Tables 2,3 and 4).

Table 2 collects all the possible combinations of big 5 characteristics that lead to an outcome. In this case, OCEAN combinations (combinations of some of the 5 characteristics) lead to high intention to post a lot on Facebook. The ‘~’ sign, indicates low values of a characteristics. For example, n*e*c combination indicates that users with high values of neuroticism, extraversion and conscientiousness are highly intended to post, as well as users with ~n*e*o combination (low values of neuroticism, and high on extraversion and openness). In specific, from our results we conclude that users with low neuroticism levels and high openness are likely to post more on Facebook and the same result (equifinality) occurs when users obtain high scores on neuroticism, extraversion and conscientiousness or low scores on neuroticism, agreeableness, extraversion and conscientiousness. The most interesting and useful outcome from this Table, as well as Tables 3 and 4 is the fact that when practitioners want to obtain high number of posts, likes or any other metric, it is enough to find users that belong to any of the combinatory solution of our model. That is, when the elevate number of posts is the final objective, analysts can choose to focus on users from diverse combinations of personality, ending up always to the same outcome. This capacity is the most powerful advantage of the equifinality provided from the present article’s results.
One of the findings that are interesting is that same result with different combinations of personality traits. This sort of fsQCA result opens up the spectrum of users with personality traits and Facebook measurable activities. The results reveal that such a relationship does exist and can be utilized to predict users' measurements in social media. Thus, the paper contributes in theory by providing empirical evidence concerning the association between Facebook activities and users' personalities in a novel way (utilizing the fsQCA method) that indicates not only the significance of this relationship but providing alternative combinations that lead to the same outcome (equifinality). Furthermore, the application of the fsQCA method to other domains is a promising future research direction. For example, in the domain of education, previous studies combine Facebook students' profiles and Big Five personality types to predict future achievements of the students (Komarraju, Karau, Schmeck, & Avdic, 2011; O'Connor & Paunonen, 2007). Applying the fsQCA method could provide valuable extensions to extant research in the domain of elections voting where it has been suggested that combining Facebook activities with personality traits can reveal future voting tendencies (Vecchione et al., 2011).

**LIMITATIONS:**
The research presented on this paper has certain limitations, such as the small sample size. However, such small sample sizes are not uncommon in Facebook personality studies and previous researches have relied upon analogous sample sizes (e.g. Amichai-Hamburger and Vinitzky (2010) and Gosling et al. (2011)). Our sample consists of a 93% of graduated or nearly graduated users. All participants are members of the same Facebook group. There is need for generalization of the findings in all educational levels and also users who do not necessary are digital friends with the authors. This last limitation though, for the moment it is inevitable, since Netvizz provides information only for users who belong to a group which the researcher has administrator permissions.

**REFERENCES**

5 DISCUSSION - CONCLUSION
One of the findings that are interesting is that same combinations are found in all three tables. That means that users with specific personalities are more likely to present elevate number of all three metrics in consideration, that is number of likes and comments posts and number of posts. From a point of view this can be also considered normal, since all three measurable activities are quite similar or relevant. Statistically speaking, the results of the research are all acceptable and indicate significant outcomes.

Managerial and theoretical implications:
The fsQCA method provides the advantage of obtaining the same result with different combinations of personality traits. This sort of fsQCA result opens up the spectrum of users with different personalities that a company should target to diffuse its marketing message posted in social media in order to achieve higher results. In the theoretical field, researchers are facilitated by basing on the present results, adding value to the existing studies and by applying our research on specific fields of study, as Tourism. Education, eHealth or other.

| Table 4: Intermediate solution indicating high intention to receive comments on a post. |
|-----------------------------------------------|------------------|------------------|
| OCEAN => PC                                    | Frequency cutoff: 2.000000 |
| consistency cutoff: 0.781850                  |                   |
| **Personalities' combinations**                | Raw coverage     | Unique coverage  |
| 1. -n^a*e^o                                   | 0.184967          | 0.057664         |
| 2. -n^e*c                                      | 0.307296          | 0.156227         |
| 3. n^e*c                                      | 0.241894          | 0.037215         |
| 4. a^e*c^o                                     | 0.126934          | 0.020449         |
| 5. -n*a^e*c                                    | 0.133014          | 0.024503         |
| 6. n^a*e^o                                    | 0.142962          | 0.034267         |
| 7. n^a*c                                        | 0.151437          | 0.035004         |

| solution coverage: 0.673729       | solution consistency: 0.809250 |

Raw coverage measures the proportion of memberships in the outcome, explained by each term of the combination (Ragin & Fiss, 2008). By this we indicate that the higher this number is, the more cases from the data set match with this combination. Unique coverage indicates the proportion of the cases from the data set that can be explained by the combination of the line. Consistency is defined as the degree to which membership in each combination term is a subset of the outcome. Solution coverage is the total raw coverage for the set of all combinations and measures the memberships in the outcome explained not by each term but from all combinations. Solution Consistency measures the degree to which membership in the whole set of the combinations is a subset of membership in the outcome (Ragin & Fiss, 2008). For the frequency, a cut-off point is set in order to ensure a minimum number of observations. Considering that our sample is rather small, we set the cut-off point in two. Ragin (2006) set the minimum threshold for consistency at 0.75. According to this research, all our results present higher consistency.


Yilmaz, C., Varnali, K., & Kasnakoglu, B. T. How do firms benefit from customer complaints? Journal of Business Research. doi: http://dx.doi.org/10.1016/j.jbusres.2015.08.038

